

WHAT IS CLAIMED IS:

1. A method for efficiently performing matrix multiplication in a vector processing computer system by using partial products that is bit-by-bit compatible with conventional matrix multiplication techniques, wherein two
5 matrices are stored in vector registers within the computer system, comprising the steps of:

storing values within a first group of vector registers within a vector processor that each comprise multiple copies of a respective individual value of a first matrix to be multiplied;

10 storing values within a second group of vector registers within a vector processor that each comprise values from a respective row of a second matrix to be multiplied;

dot multiplying each of said first vector registers having values from a single row of said first matrix by each of said second vector registers;

15 adding the values obtained by said step of multiplying to a third group of vector registers corresponding to a product matrix; and

repeating said steps of multiplying and adding, for said first vector registers having values from every row of said first matrix.

2. The method of claim 1, wherein said step of multiplying is carried out
20 in a manner such that vector registers containing multiple copies of each value from a single row are each multiplied by vector registers from a second group of vector registers from row numbers corresponding to the element number from which the elements contained within the first group of vector registers is taken.

3. The method of claim 1, wherein said step of adding is carried out such that the values obtained by dot multiplying vector registers from the first group of vector registers having values from a single row are added to the third group of vector registers corresponding to the product matrix in such a manner that the
5 vector registers obtained correspond to the same row of the product matrix as the row from which their values were taken from said first matrix.

4. A system for efficiently performing matrix multiplication in a vector processing computer system by using partial products that produces a bit-by-bit compatible result with conventional matrix multiplication techniques, wherein
10 values of two matrices are stored in vector registers within the computer system, comprising the steps of:

means for storing values within a first group of vector registers within a vector processor that each comprise multiple copies of each individual value of a first matrix to be multiplied;

15 means for storing values within a second group of vector registers within a vector processor that each comprise values from each row of a second matrix to be multiplied;

means for dot multiplying each of said first vector registers having values from a single row of said first matrix by each of said second vector registers; and

20 means for adding the values obtained by said step of multiplying to a third group of vector registers corresponding to a product matrix.

5. The system of claim 4, wherein said means for dot multiplying is configured to perform the multiplication in a manner such that vector registers containing multiple copies of each value from a single row are each multiplied by
25 vector registers from a second group of vector registers from row numbers

03846678-032401
TOTAL 866T860

corresponding to the element number from which the elements contained within the first group of vector registers is taken.

6. The system of claim 4, wherein said means for adding is configured to multiply in a manner such that the values obtained by dot multiplying vector registers from the first group of vector registers having values from a single row are added to the third group of vector registers corresponding to the product matrix in such a manner that the vector registers obtained correspond to the same row of the product matrix as the row from which their values were taken from said matrix.

7. A method for multiplying two matrices that is bit-by-bit compatible with conventional matrix multiplication techniques in a vector processor, comprising the steps of:

(i) replicating the value of each element of a first matrix to form corresponding vectors each having a number of elements corresponding to the number of columns in a second matrix, and storing said vectors in a first set of registers;

(ii) storing the value of the elements of each row of said second matrix as corresponding vectors in a second set of registers;

(iii) multiplying a vector in one of said first set of registers with a vector in one of said second set of registers such that the vector in one of the first set of registers multiplied with a vector in one of the second set of registers is selected from a single row within the first matrix, and adding the resulting products as a vector in one of a set of third registers corresponding to the rows of a product matrix equivalent to the row from which the element stored in the vector of the first set of registers is taken; and

(iv) iteratively repeating step (iii) for each vector in said first set of registers.

8. The method of claim 7, wherein step (iii) comprises multiplying a vector whose values are elements contained in a given row of the first matrix with
5 a vector containing values of elements contained in the corresponding row of the second matrix.

9. The method of claim 8, wherein the resulting product is stored in one of said third set of registers which corresponds to said given row in a third matrix.

10. A method for efficiently calculating the product of a first matrix and a second matrix in a vector processing system that is bit-by-bit compatible with conventional matrix multiplication techniques, wherein said first matrix is stored in a plurality of first vector registers, and wherein said second matrix is stored in a plurality of second vector registers, and wherein each of said first plurality of
15 vector registers corresponds to an individual row of said first matrix, and each of said second plurality of vector registers corresponds to an individual row of said second matrix, comprising the steps of:

- (1) selecting a vector register corresponding to a row of said first matrix;
- (2) storing multiple copies of a single value from the register selected in
20 step 1 in a vector register and wherein the single value selected from the vector register selected in step 1 has a specific index within the vector register selected in step 1;
- (3) calculating the dot product of the vector register containing multiple copies of a single value that was stored in step 2 and the vector register containing

elements from the row number of the second matrix which corresponds to the index of the element stored within the vector register in step 2;

(4) adding the dot product calculated in step 3 in one of a third group of vector registers that form a third matrix, which is the product of the first two
5 matrices, by adding the dot product value in the register corresponding to the row of the product matrix that is equivalent to the row of the first matrix that corresponds to the vector register selected in step 1;

(5) repeating steps 2 through 4 for each of the elements contained within the register selected in step 1; and

10 (6) repeating steps 1 through 5 for each of the registers corresponding to each row of the first matrix.

11. The method of claim 10, wherein said step of storing multiple copies of a single value from the register selected in step 1 further comprises storing the multiple copies in a vector register having the same number of elements as the
15 number of elements of each row of said second matrix.

12. The method of claim 10, wherein said third group of vector registers corresponding to a third matrix, or product matrix, is initialized by storing initial values of each element within the third group of vector registers.

20 13. The method of claim 12, wherein said initialization occurs by storing the value zero in each location of each vector of said third group of vector registers.

25 14. A method for efficiently calculating the product of a first matrix and a second matrix in a vector processing system that is bit-by-bit compatible with conventional matrix multiplication techniques, wherein said first matrix is stored

in a plurality of first vector registers, and wherein said second matrix is stored in a plurality of second vector registers, comprising the steps of:

(i) dividing each of said first and second matrices into a plurality of submatrices;

5 (ii) assigning values of each submatrix associated with said first and second matrices to said pluralities of first and second vector registers, respectively, in a manner such that each of said first plurality of vector registers corresponds to an individual row of a submatrix of said first matrix, and each of said second plurality of vector registers corresponds to an individual row of a submatrix of said second
10 matrix;

(iii) selecting a vector register corresponding to a row of a submatrix of said first matrix;

(iv) storing multiple copies of a single value from the register selected in step (iii) in a vector register, and wherein the single value selected from the vector
15 register selected in step (iii) has a specific index within the vector register selected in step (iii);

(v) calculating the dot product of the vector register containing multiple copies of a single value that was stored in step (iv) and the vector register containing elements from the row number of each submatrix of the second matrix
20 which corresponds to the index of the element stored within the vector register in step (iv);

(vi) adding the dot product calculated in step (v) to one of a third group of vector registers that form a third matrix, which is the product of the first two matrices, by adding the dot product value in the register corresponding to the row
25 of the submatrix within the product matrix that is equivalent to the row of the first matrix that corresponds to the vector register selected in step (iii);

(vii) repeating steps (iv- vi) for each of the elements contained within the register selected in step (i); and

(viii) repeating steps (iii-vii) for each of the registers corresponding to each row of each submatrix of the first matrix.

5 15. The method of claim 14 wherein said step of storing multiple copies of the single value from the register selected in step (iii) in a vector register further comprises storing the multiple copies in a vector register having the same number of elements as the number of elements of each row of each submatrix of the second matrix with which it is multiplied.

10 16. The method of claim 14 wherein said third group of vector registers corresponding to a third matrix, or product matrix, is initialized by storing the initial values of each element within the third group of vector registers.

15 17. The method of claim 16 wherein said initialization occurs by storing the value zero in each location of each vector of said third group of vector registers.

15 18. The method of claim 14 wherein said steps of calculating and adding are performed simultaneously by way of an integrated vector processing function.